

2850-106.ST25.txt
SEQUENCE LISTING

<110> Yokoyama, Shigeyuki
Shirouzu, Mikako
Sakamoto, Ayako
Sakamoto, Kensaku

<120> NON-NATURALLY-OCCURRING AMINO ACID-CONTAINING PROTEIN EXPRESSION METHOD

<130> P/2850-106

<140> 10/532,948
<141> 2005-11-10

<150> PCT/JP03/14028
<151> 2003-10-31

<150> 2002-318846
<151> 2002-10-31

<160> 32

<170> PatentIn version 3.3

<210> 1
<211> 167
<212> DNA
<213> Artificial

<220>
<223> an artificial base sequence consisting of a leader sequence of human tRNA gene, and the tRNA^{tyr} gene of B. stearothermophilus with a CUA anticodon, but without the terminal CCA sequence and a transcription terminator

<400> 1
agcgctccgg tttttctgtg ctgaacctca ggggacgccg acacacgtac acgtcggagg 60
ggtagcgaag tggctaaacg cggcggactc taaatccgct ccctttgggt tcggcggttc 120
gaatccgtcc ccctccagac aagtgcggtt tttttctcca gctccccg 167

<210> 2
<211> 44
<212> DNA
<213> Artificial

<220>
<223> a PCR primer used for amplifying a part of genomic DNA of E. coli

<400> 2
ggaattccat atggcaagca gtaacttgat taaacaattg caag 44

<210> 3
<211> 50
<212> DNA
<213> Artificial

<220>
<223> a PCR primer used for amplifying a part of genomic DNA of E.

coli.

<400> 3
gccgaagctt gtcgactttc cagcaaatca gacagtaatt ctttttaccg 50

<210> 4
<211> 40
<212> DNA
<213> Artificial

<220>
<223> a PCR primer used in the overlapping extension in the present invention

<400> 4
aggatcgaag ccgcaagcga gcgcatcg gccttgcgcc 40

<210> 5
<211> 40
<212> DNA
<213> Artificial

<220>
<223> a PCR primer used in the overlapping extension in the present invention

<220>
<221> misc_feature
<222> (16)..(16)
<223> m represents c or a

<220>
<221> misc_feature
<222> (17)..(18)
<223> n is a, c, g, or t

<400> 5
aggatcgaag ccgcamnga gcgcatcg gccttgcgcc 40

<210> 6
<211> 33
<212> DNA
<213> Artificial

<220>
<223> a PCR primer used in the overlapping extension in the present invention

<400> 6
acgggtgtgt gctgtctatt ggtggttctg acc 33

<210> 7
<211> 33
<212> DNA
<213> Artificial

<220>
<223> a PCR primer used in the overlapping extension in the present
Page 2

invention

<400> 7
acggtgtggt gctggcaatt ggtggttctg acc 33

<210> 8
<211> 33
<212> DNA
<213> Artificial

<220>
<223> a PCR primer used in the overlapping extension in the present invention

<400> 8
acggtgtggt gctgaacatt ggtggttctg acc 33

<210> 9
<211> 33
<212> DNA
<213> Artificial

<220>
<223> a PCR primer used in the overlapping extension in the present invention

<400> 9
acggtgtggt gctgtgcatt ggtggttctg acc 33

<210> 10
<211> 32
<212> DNA
<213> Artificial

<220>
<223> a PCR primer used in the overlapping extension in the present invention

<400> 10
ttcttcggat ccaaccagac tgcgccgcct tc 32

<210> 11
<211> 30
<212> DNA
<213> Artificial

<220>
<223> a PCR primer used in the overlapping extension in the present invention

<400> 11
gatcatctgg ttaacggaga agtgtttgcc 30

<210> 12
<211> 26
<212> DNA
<213> Artificial

<220>
 <223> a PCR primer used in the overlapping extension in the present invention

<400> 12
 gaccttcctg tgcgatattg gcaaac 26

<210> 13
 <211> 12
 <212> DNA
 <213> Artificial

<220>
 <223> the box A consensus sequence

<220>
 <221> misc_feature
 <222> (2)..(2)
 <223> r represents g or a

<220>
 <221> misc_feature
 <222> (5)..(6)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (10)..(10)
 <223> n is a, c, g, or t

<400> 13
 trgcnnagyn gg 12

<210> 14
 <211> 11
 <212> DNA
 <213> Artificial

<220>
 <223> the box B consensus sequence

<220>
 <221> misc_feature
 <222> (8)..(8)
 <223> n is a, c, g, or t

<400> 14
 ggttcgantc c 11

<210> 15
 <211> 20
 <212> DNA
 <213> Artificial

<220>
 <223> a sequence of a primer binding site pbs1

<400> 15

2850-106.ST25.txt

agcgagtgtt aaccctgcct 20

<210> 16
 <211> 20
 <212> DNA
 <213> Artificial

<220>
 <223> a sequence of a primer binding site pbs2

<400> 16
 cgactacgat attcgcgcag 20

<210> 17
 <211> 12
 <212> DNA
 <213> Artificial

<220>
 <223> a sequence of a BstXI-1 site

<400> 17
 ccagcagact gg 12

<210> 18
 <211> 12
 <212> DNA
 <213> Artificial

<220>
 <223> a sequence of a BstXI-2 site

<400> 18
 ccagcttcct gg 12

<210> 19
 <211> 63
 <212> DNA
 <213> Artificial

<220>
 <223> a nucleotide sequence coding a short peptide used for
 substitution of green fluorescent protein (cyanfluorescent
 mutation)

<400> 19
 atgggaacta gtccatagtg gtggaattct gcagatatcc agcacagtgg cggccgccgc 60
 gtc 63

<210> 20
 <211> 11
 <212> DNA
 <213> Artificial

<220>
 <223> another box B consensus sequence

```

<220>
<221> misc_feature
<222> (8)..(8)
<223> n is a, c, g, or t

<400> 20
agttcgantc t 11

<210> 21
<211> 31
<212> DNA
<213> Artificial

<220>
<223> a sequence of a primer used for amplifying the sequence of SEQ ID
No. 1

<400> 21
cacagaattc tcgggagctg gagaaaaaaa c 31

<210> 22
<211> 30
<212> DNA
<213> Artificial

<220>
<223> a sequence of another primer used for amplifying the sequence of
SEQ ID No. 1

<400> 22
cacaaagctt agcgctccgg tttttctgtg 30

<210> 23
<211> 40
<212> DNA
<213> Artificial

<220>
<223> a sequence of a primer set used for amplifying a fragment having
a primer binding site pbs1 upstream of the sequence of SEQ ID No.
1 and BstXI-1 site downstream thereof

<400> 23
agcgagtgtt aaccctgcct agcgctccgg tttttctgtg 40

<210> 24
<211> 38
<212> DNA
<213> Artificial

<220>
<223> a sequence of a primer set used for amplifying a fragment having
a primer binding site pbs1 upstream of the sequence of SEQ ID No.
1 and BstXI-1 site downstream thereof

<400> 24
acacacccag cagactggcg ggagctggag aaaaaaac 38

```

2850-106.ST25.txt

<210> 25
 <211> 38
 <212> DNA
 <213> Artificial

 <220>
 <223> a sequence of a primer set used for amplifying a fragment having
 a BstXI-1 site upstream of the sequence of SEQ ID No. 1 and
 another BstXI-1 site downstream from the first BstXI-1 site

 <400> 25
 acacacccag cagactggag cgctccggtt tttctgtg 38

 <210> 26
 <211> 38
 <212> DNA
 <213> Artificial

 <220>
 <223> a sequence of a primer set used for amplifying a fragment having
 a BstXI-1 site upstream of the sequence of SEQ ID No. 1 and
 another BstXI-1 site downstream from the first BstXI-1 site

 <400> 26
 acacacccag cttcctggcg ggagctggag aaaaaaac 38

 <210> 27
 <211> 38
 <212> DNA
 <213> Artificial

 <220>
 <223> a sequence of a primer set used for amplifying a fragment having
 a BstXI-2 site upstream of the sequence of SEQ ID No. 1 and a
 primer binding site pbs-2

 <400> 27
 acacacccag cttcctggag cgctccggtt tttctgtg 38

 <210> 28
 <211> 40
 <212> DNA
 <213> Artificial

 <220>
 <223> a sequence of a primer set used for amplifying a fragment hving a
 BstXI-2 site upstream of the sequence of SEQ ID No. 1 and a
 primer binding site pbs-2

 <400> 28
 ctgcgcgaat atcgtagtcg cgggagctgg agaaaaaac 40

 <210> 29
 <211> 424
 <212> PRT
 <213> Escherichia coli

 <400> 29

2850-106.ST25.txt

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30

Pro Ile Ala Leu Tyr Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly
65 70 75 80

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Asp Phe Ala Cys Leu Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

2850-106.ST25.txt

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210> 30

<211> 135

<212> DNA

<213> Artificial

<220>

<223> a sequence of one of the induced expression systems prepared in
Example 2 (TetBst0)

<400> 30

tctccctatc agtgatagag atcggagggg tagcgaagtg gctaaacgcg gcggactcta 60

aatccgctcc ctttggttc ggcggttcga atccgtcccc ctccagacaa gtgcggtttt 120

tttctccagc tcccg 135

<210> 31
 <211> 145
 <212> DNA
 <213> Artificial

<220>
 <223> a sequence of one of the induced expression systems prepared in
 Example 1 (TetBst1)

<400> 31
 tctccctatc agtgatagag atccgtacac gtcggagggg tagcgaagtg gctaaacgcg 60
 gcggactcta aatccgctcc ctttgggttc ggcggttcga atccgtcccc ctccagacaa 120
 gtgcggtttt tttctccagc tcccg 145

<210> 32
 <211> 155
 <212> DNA
 <213> Artificial

<220>
 <223> a sequence of one of the induced expression systems prepared in
 Example 2 (TetBst2)

<400> 32
 tctccctatc agtgatagag atccgccgac acacgtacac gtcggagggg tagcgaagtg 60
 gctaaacgcg gcggactcta aatccgctcc ctttgggttc ggcggttcga atccgtcccc 120
 ctccagacaa gtgcggtttt tttctccagc tcccg 155